DOCUMENT RESUME

ED 355 285 TM 019 642

AUTHOR Hume, Jerry
INSTITUTION National Assessment Governing Board, Washington, DC.
TITLE Future of NAEP.
PUB DATE 19 Nov 92
NOTE 35p.
PUB TYPE Reports - Evaluative/Feasibility (142)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Achievement Tests; Advisory Committees; Comparative Analysis; *Data Collection; *Educational Assessment; Educational Policy; Elementary Secondary Education; Evaluation Utilization; Federal Legislation; *Futures (of Society); Guidelines; Longitudinal Studies; National Competency Tests; *National Surveys; Predictive Measurement; School Districts; State Programs; *Test Use
IDENTIFIERS Focus Groups Approach; *National Assessment of Educational Progress

ABSTRACT

This memorandum presents the report and recommendations of the Ad Hoc Committee on the Future of the National Assessment of Educational Progress (NAEP). The Ad Hoc Committee was created by the National Assessment Governing Board in 1992 to prepare recommendations on non-mandated subjects to be assessed by the NAEP and to prepare a discussion paper on policy issues facing the NAEP. The Committee received 107 written responses to the discussion paper from interested individuals and agencies. Focus groups drawn from similar entities provided further commentary. Unanimity of opinion did not exist among commentators, but the following views were predominant: (1) the NAEP should collect state-level data regularly; (2) the NAEP should reflect current and evolving practices; (3) the NAEP should be independent; (4) achievement levels should be set and used; (5) when possible, the NAEP should be equated with tests from other nations; (6) state and local districts should use the NAEP to link local results with national and international results; (7) Congress should allow states and districts to use NAEP results; and (8) legislation should be amended to permit an annual assessment. Recommendations based on these views are presented. Attachments include a schedule of the NAEP assessment through the year 2000, analysis of discussion paper comments, and alternative scenarios for the future of the NAEP. (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED355285

FUTURE OF NAEP

National Assessment Governing Board

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

R. FIELDS

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)"

TM019642

BEST COPY AVAILABLE

TM

THE NATION'S REPORT CARD

National Assessment Governing Board

National Assessment of Educational Progress

TO : Executive Committee

FROM : Jerry Hume
Chair
Ad Hoc Committee on the Future of NAEP

SUBJECT : Report and Recommendations

DATE : November 19, 1992

The purpose of this memorandum is to transmit the report and recommendations of the Ad Hoc Committee on the Future of NAEP.

A. BACKGROUND

On March 6, 1992, the Ad Hoc Committee on the Future of NAEP was created by the National Assessment Governing Board. The charge to the committee was to prepare recommendations on non-mandated subjects to be assessed by the National Assessment and to prepare recommendations for a document to be referred to as "Positions on the Future of NAEP, Second Edition, 1992" for presentation at the November 1992 Board meeting.

On August 7, 1992, the Governing Board unanimously adopted the recommendation of the Committee regarding non-mandated subjects to assess; the Governing Board also approved for dissemination and public comment, a discussion paper on policy issues facing the National Assessment, prepared by the Committee.

The Committee developed a discussion paper designed to engage the public in examining issues confronting NAEP. The paper was made available upon request without charge and widely disseminated during August and September. The Committee received 107 written responses from states, public and private schools, education associations and interest groups, higher education, the measurement community, business, individuals, other government agencies, and educational technical assistance organizations.

In addition, two day-long focus groups were conducted to examine the issues in depth. The first was conducted with the Education Information Advisory Council Task Force on Assessment, an arm of the Council of Chief State School Officers comprised of state assessment directors. The second was conducted as the Forum on the Future of NAEP. The Forum included a diverse group of individuals, with participation from civil rights groups, public and private school education associations, state education agencies, business, test publishers, the National Education Goals Panel, the National Governors Association, congressional staff, school districts, and the measurement community.

800 North Capitol Street, N.W.
Suite 825
Mailstop 7583
Washington, D.C. 20002-4233
(202) 357-6938

TM019642

B. SUMMARY ANALYSIS

The attachment entitled "Analysis of Comments on the Future of NAEP Discussion Paper" (Tab 2) describes in detail the results of the committee's efforts to obtain external advice. While there is not unanimity of opinion among the commenters, there is a clearly predominant view:

1. NAEP should regularly collect state-level data.

2. NAEP should reflect both current and evolving instructional practice. It should neither be determined by nor ignore voluntary national content and performance standards as they are developed. Instead, a balance should be achieved, through the national consensus process used in developing assessment frameworks, to align appropriately NAEP with the standards through successive administrations of a subject area assessment.

3. NAEP should be independent of any entity created to review and/or certify standards and assessments.

4. Achievement levels have the potential for improving public understanding of National Assessment results. NAEP should continue to report results using achievement levels and continue improving the process by which they are set.

5. It would be useful to know how students in other nations perform on NAEP or NAEP items as a reference point for understanding U.S. student performance; this includes equating NAEP results with international studies where possible. However, technical issues must be satisfactorily addressed, nations must be found that would be willing to participate and NAEP's mission must remain that of a domestic assessment.

6. States and local districts should be permitted to use NAEP to link the results of their local assessments with national and international results. Participation must be voluntary and appropriate controls must be installed to assure the continuing credibility and integrity of NAEP. Using NAEP in this way will pose technical challenges; thus, research and development in this area should continue.

7. Congress should allow States and districts to use NAEP test items, data, instruments and analytic mechanisms, at their discretion and cost, according to NAEP requirements, and subject to federal oversight, to report district and school (but not individual student) results.

8. The National Assessment legislation should be amended to permit annual assessments.

9. Additional Issues: Several respondents identified issues that are not addressed in the discussion paper on the Future of NAEP.

a. Testing Linguistic Minorities

Several respondents and a Forum participant raised the concern that the significant growth in the numbers of limited English proficient students nationally, and the large proportion of such students in certain states, poses technical and policy matters that need to be addressed by the National Assessment. For example, one respondent estimated that "...by the year 2000, the majority of students in California may well be nonnative [sic] English speakers." The question is--how representative can state/national NAEP results be if such students are not included in the assessment and how can such students be fairly assessed?

b. Participation of Disabled Students

Two respondents raised the concern that exclusion rates from NAEP of disabled students are very high and too variable among the states. The point made by the respondents is that high and variable unexplained exclusions of disabled students impact on the comparability of state assessments and the comprehensiveness of national assessments.

c. Linking Curricular Practices and Policies to Outcomes

Two respondents addressed the potential in NAEP, through the background questionnaires, to do a better job of collecting and analyzing information on the conditions of instruction in the nation's classrooms and of linking this with information on student outcomes. One of these respondents suggested using NAEP as a mechanism to define and collect information related to the educational delivery standards discussed in the report of the National Council on Education Standards and Tests.

C. A LOOK INTO THE FUTURE

In order to put these perspectives into a context of the future, several scenarios were developed, based on what is known about current proposals and on judgments of the plausible responses of affected agencies. The scenarios are described in the attachment entitled "Alternative Future Scenarios." Four scenarios are described:

- full implementation of the NCEST proposal
- content standards developed in the absence of a certification process
- adoption of school delivery standards
- rapid state response to national standards

Clearly, these scenarios are not exhaustive, and they are speculative. But they do seem to point in a certain direction for NAEP.

The scenarios indicate that, even under the most optimistic projections, the NCEST proposal is not likely to be fully implemented before 1998. It is not likely that mechanisms will be in place such that individual student tests could be aggregated to provide comparable school, district, state and national results much before the end of the decade, although it is possible that work to develop those mechanisms will be underway and widely supported. The possibility exists that, sometime in the future, assuming standards are widely implemented and comparability among assessment systems is attained, NAEP in its present configuration may not be needed, especially if the assessment systems contain controls that monitor and protect against abuse. However, it is not yet clear that sufficient incentives, resources or political will exist to achieve an integrated national system of assessments.

D. CONCLUSIONS

Until such time as an integrated national system of assessments comes into existence, NAEP is in a unique position to serve as a monitor of national education progress and provide state-comparable data on education performance and conditions of instruction. Most of the respondents to the discussion paper issues believe that state-level reporting of NAEP is important. They believe it is useful for monitoring state education progress and for informing state-level policy development. They believe that state-level reporting of results does not transform NAEP into an accountability instrument.

Whether it should regularly provide state-comparable data is the central policy question for NAEP at present. Conference report language for a bill that failed in the closing hours of the 102nd Congress indicates that there was general support for continuing state trial assessments, at least into 1994. Had it been enacted, the NAEP provisions still represent what would have been the continuation of an experiment, not the adoption of policy. However, the response to the discussion paper indicates a strong desire on the part of a divergent group of interests for a regular state NAEP component, particularly on the part of states.

Another pattern that emerges from the analysis is that the period during which standards and assessment systems are developed will and should be a time of research, development and evaluation, and that a program of research and development for NAEP is desirable. This is a theme that arises in the responses on achievement levels, using NAEP as an anchor, lifting the prohibition, and an international component to NAEP, as well as the additional issues relating to linguistic minorities, disabled students, and opportunity to learn. The general view expressed by the respondents is that NAEP items, instruments and analytic tools are excellent resources that should be appropriately protected, but also appropriately developed and applied in developing a national system of assessments.

E. RECOMMENDATIONS

The Committee recommends the following:

1. NAEP should regularly collect state-level data.

2. NAEP should reflect both current and evolving instructional practice. It should neither be determined by nor ignore voluntary national content and performance standards as they are developed. Instead, a balance should be achieved, through the national consensus process used in developing assessment frameworks, to align appropriately NAEP with the standards through successive administrations of a subject area assessment.

3. The Committee believes that taking a position at this time on the relation between NAEP and any entity created to review and/or certify standards and assessments is premature. The Committee recognizes that current law provides mechanisms designed to ensure the independence of NAEP, to keep it free from inappropriate influences and special interests, and to assure its integrity and credibility. The Committee believes that these should continue as guiding principles for NAEP, but that it is not appropriate to prescribe a relationship with an entity that does not yet exist.

4. Achievement levels have the potential for improving public understanding of National Assessment results. NAEP should continue to report results using achievement levels and continue improving the process by which they are set.

5. It would be useful to know how students in other nations perform on NAEP or NAEP items as a reference point for understanding U.S. student performance; this includes equating NAEP results with international studies where possible. However, technical issues must be satisfactorily addressed, nations must be found that would be willing to participate and NAEP's mission must remain that of a domestic assessment.

6. States and local districts should be encouraged and assisted in using NAEP to link the results of their local assessments with national and international results. Participation must be voluntary and appropriate controls must be installed to assure the continuing credibility and integrity of NAEP. Using NAEP in this way will pose technical challenges; thus, research and development in this area should continue.

7. Congress should allow States and districts to use NAEP test items, data, instruments and analytic mechanisms, at their discretion and cost, according to NAEP requirements, and subject to federal oversight, to report district and school (but not individual student) results.

8. The National Assessment legislation should be amended to permit annual assessments.

9. Concerns about participation in NAEP of students with limited English proficiency (i.e. linguistic minorities) and students with disabilities should be addressed through follow-up contacts with the respondents and with a program of research.

10. Work should continue, through appropriate committees of the Board, to examine ways to improve NAEP's capacity to report on the relationships between educational practices and outcomes.

11. The Governing Board should adopt the schedule of assessments through the year 2000 as it appears in Tab 1.

12. The Governing Board should reaffirm its policy that the federal government should pay the full cost of state-level NAEP assessments.

Attachments:

Tab 1 -- Schedule of NAEP Assessments
Tab 2 -- Analysis of Comments
Tab 3 -- Alternative Future Scenarios

TAB 1

National Assessment of Educational Progress
Schedule of Subjects to be Assessed: 1994-2000

Assumptions:
1. All assessments are national and state-level.
2. The Third International Mathematics and Science Study (TIMSS) will be conducted in 1995 and 1999.
3. Assessment frameworks will be developed in civics, economics, foreign language and world history.
4. Congress provides authority for state assessments and annual data collection and reporting.

1994 -	Reading
	Geography
	U.S. History
1995 -	Mathematics
	Science
	Writing
1996 -	Reading
	Arts
	Civics
1997 -	Mathematics
	Economics
	Foreign Language
1998 -	Reading
	Geography
	U.S. History/World History
1999 -	Mathematics
	Science
	Writing
2000 -	Reading
	Civics
	Optional

Analysis
The frequency of assessments is consistent with current law:
1. Reading and mathematics assessments are conducted biennially in alternating years. Current law requires biennial assessments of reading and mathematics.
2. Science assessments are conducted once every four years,[1] as under current law, and are always paired with mathematics. They are scheduled in 1995 and 1999 to facilitate linking TIMSS and NAEP results and to consider international results in setting achievement levels.
3. All non-mandated subjects recommended for assessment by the Board are included in the schedule. Only civics is assessed twice before the end of the decade, although an "optional" slot is provided in the year 2000.
4. History and Geography are paired. Assessments are conducted on a once-per-four-year cycle; current law requires assessments "at least once every six years."
5. Writing is on a once-per-four-year cycle, as under current law.[2]

---

[1] Except for a 5 year interval from 1990 in order to align with TIMSS.

[2] Except for a 3 year interval from 1992.

TAB 2

THE NATION'S REPORT CARD

National Assessment Governing Board

National Assessment of Educational Progress

Analysis of Comments on The Discussion Paper on the Future of NAEP

Background

On March 6, 1992, the Ad Hoc Committee on the Future of NAEP was created by the National Assessment Governing Board. The charge to the committee was to prepare recommendations on non-mandated subjects to be assessed by the National Assessment and to prepare a document to be referred to as "Positions on the Future of NAEP, Second Edition, 1992."

On August 7, 1992, the Governing Board unanimously adopted the recommendation of the Committee regarding non-mandated subjects to assess; the Governing Board also approved for dissemination and public comment, a discussion paper on policy issues facing the National Assessment, prepared by the Committee.

On August 22, the Governing Board published a notice in the Federal Register inviting public comment on the issues in the discussion paper. Education Daily and Report on Education Research also ran stories about the paper and the invitation to comment.

Approximately 4,500 copies of the discussion paper have been distributed to parents, civil rights and education advocacy groups, individuals and organizations representing the education measurement and research community, schools, school administrators, business organizations, state education agencies and governors. In all, we have received 107 written responses. We also have discussed the paper at a national meeting of deans of education colleges.

In addition, we have conducted two day-long "focus groups" to examine the issues in the discussion paper in depth. One focus group was with the Education Information Advisory Committee (EIAC) of the Council of Chief State School Officers. As a result, EIAC passed a resolution that states its recommendations on the issues discussed in the paper (attached).

The second focus group was conducted as the Forum on the Future of NAEP. The Forum participants comprised a diverse group of individuals knowledgeable about NAEP. The range of interests represented by the individuals present included: civil rights groups, public and private school education associations, state education agencies, business, test publishers, the National Education Goals Panel, the National Governors Association, congressional staff, school districts, and the measurement community. The intent of this focus group was to collect the

800 North Capitol Street, N.W.
Suite 825
Mailstop 7583
Washington, D.C. 20002-4233
(202) 357-6938

individual opinions of the participants and to determine where there was agreement and where there was disagreement. We stipulated that our report would focus on the substance of the discussion and would not attribute to any individual or organization endorsement of any position without their consent.

Analysis

Written responses were received from 107 individuals and organizations. The category and number of responses are displayed in Table 1.

Table 1.

Associations	22
Schools/Districts	21
Higher Education	18
Measurement	12
States	14
Individuals	9
Business	6
Technical Assistance Agencies	3
Other Government	2
	107

Not all issues raised in the discussion paper were addressed by every respondent and, in some cases, the comments provided were not germane. In addition, care was taken in the analysis of the comments to attribute a position to the respondent only where one was clearly stated; silence on a matter was considered a non-response. Therefore, the number of responses to an issue is somewhat fewer than the total number of responses received. A few respondents raised issues not discussed in the paper; these issues will be discussed in the analysis below. In weighing the opinions that have been expressed, it should be noted that the sample is self-selected, and thus not necessarily representative.

The response rate for each issue varied considerably, from 25% to 68%. Only three issues had a response rate approaching or exceeding 50%--number 1., on state NAEP (68%), number 2., on alignment with national standards (50%), and number 7., on removing the prohibition (48%). Response rates to the other issues ranged from 25% to 39%.

Of those responding, 82% were in support of regular state NAEP, 74% felt that NAEP should strike a balance between new content standards and current (and evolving) practice, and 67% felt that the prohibition should be relaxed.

The comments on each issue are discussed in detail below.

1. Role and Purpose of the National Assessment

The fundamental question was whether NAEP regularly should collect state-level data. This question was posed in the context of the decision of the National Education Goals Panel to use NAEP as a primary source of data for reporting progress on Goal 3 and the hypothesis that this could transform NAEP into an accountability instrument. NAEP's future role, therefore, is cast in a dichotomy between being a monitor or an accountability instrument.

The report of the National Council on Education Standards and Tests views NAEP's role as a "program monitor," but does not define the term. Some commentators have suggested that the mere reporting of state-level data, with attendant state rankings, necessarily transforms NAEP into an accountability instrument.

While some respondents agreed with the latter proposition, most did not. To the extent that accountability involves an obligation of an entity to report its performance to a superordinate entity that can exercise direct consequences tied to the report of performance, the predominant view is that state-level reporting of NAEP results does not meet this definition of accountability. Participation in state NAEP is voluntary, not obligatory. In state NAEP, the data are provided to the state for use as it sees fit, there is no superordinate entity positioned to take direct consequences. Finally, any consequences ensuing from state-level reporting are subject to the control of the state; there are no superordinate bodies that can apply consequences as a result of NAEP state-level data. Thus, the predominant view is that there is very little to fear that state-level reporting will transform NAEP into an accountability instrument.

Table 2.

Should NAEP Regularly Collect State-level Data?

	Total	yes	no
Schools/Districts	21	14	3
States	14	11	0
Associations	22	12	2
Higher Education	18	9	1
Measurement	12	8	3
Business	6	4	0
Individuals	9	1	3
Other Government	2	1	0
Technical Assist	3	0	01
	107	60	13

Across the range of respondents, the predominant view is that NAEP should continue to collect and report State-level data. The EIAC resolution states that the "NAEP legislation should be amended...to continue to conduct State-level assessments, on a voluntary basis." The discussion of the Forum participants, likewise, indicated general acceptance of continued state-level data; some participants suggested that it continue as a trial, subject to evaluation.

Of those respondents who felt that NAEP should not collect state-level data, the reasons ranged from "...the purposes set forth for collecting state-level NAEP cannot be fulfilled [i.e. to inform instruction and policy development]..." to "...the possible advantages of full-scale state-level reporting are far outweighed by the danger that NAEP will pre-empt state-level assessment initiatives rather than merely supplement them."

However, it should be noted that the former respondent ..."strongly urge[d]..NAEP to...allocate more assessment time to an expanded set of background questions to document the conditions of instruction in our nation's classrooms..." Presumably, this suggests that NAEP achievement data coupled with information on instructional practices (e.g. degree of implementation of national content standards) would be useful in informing state-level policy in ways that could bear on instructional practice. The positions of the states and of EIAC in support of state-level NAEP indicate that they do not have serious concerns that NAEP will pre-empt state-level assessment initiatives.

2. Alignment of the National Assessment with Nationally Certified Content and Student Performance Standards
3. Assessment Frameworks

These two issues are linked. The question of whether and how the National Assessment would reflect national standards is inseparable from the question of what the assessment frameworks (i.e. the blueprint for each assessment) contain. Three fundamental values were expressed: that NAEP should "drive" reform; that NAEP should only measure the effects of reform; and that NAEP should remain insulated from education faddism. Those who hold that it should drive reform suggest that national content standards should be the sole determinant of the content of NAEP assessments. Those who believe that NAEP should measure the effects of reform, but not "drive" it suggest that national content standards cannot be ignored and are a legitimate, but not the only, consideration in the development of assessment frameworks. Those who believe that NAEP should remain insulated from faddism suggest that national standards should not be incorporated into NAEP until after they have been widely implemented and comprise representative education practice.

The predominant view of the respondents, of EIAC, and of the Forum participants, is that NAEP cannot and should not ignore national standards. This view holds that standards will be implemented in different ways and at different rates in schools throughout the nation. But, to the extent that the standards represent a broad consensus, they indicate a clear vision of shared education aspirations.

Most felt that NAEP should evolve toward the standards in successive administrations of a content area assessment. They suggested that the appropriate balance between the standards and representative practice should be accomplished through the consensus process used by the Governing Board to develop assessment frameworks for NAEP.

Table 3.
Should NAEP reflect both national standards and current (and evolving) instructional programs?

		Yes	No	
	Total		Current and evolving only	Standards only
Schools/Districts	21	8	2	1
States	14	8	1	
Associations	22	7	2	1
Higher Education	18	4	2	1
Measurement	12	8	0	
Business	6	1	0	2
Individuals	9	2	1	
Other Government	2	1	0	
Technical Assist	3	1	1	
	107	40	09	5

4. Role of the National Assessment in Relation to Organizations that May be Established to Review or Certify National Standards and a System of Assessments

Most respondents did not respond to this issue. EIAC expressed the view, echoed by others, that it is premature and highly speculative to discuss the relationships that should exist between NAEP and organizations that do not yet exist. The discussion of the Forum participants, while recognizing the speculative nature of the discussion, led to the statement of some considerations to take into account as the role is being formulated:

- Maintain NAEP's independence and credibility
- Assure that the relationship is not prescriptive
- Recognize that the relationship will operate in a "fluid" environment
- Work to see that report releases are "aligned" or coordinated

Of those respondents who provided comments, most suggested that NAEP should be independent of the new entity and be subject to its review only on a voluntary basis.

Typical of the argument given in favor of NAEP being subject to review by the new entity is the following: "If NAEP were to become the basis for a national system of assessments, the public should have the assurance that the assessment meets professional standards for content coverage and technical rigor."

Table 4.

Should the National Assessment be subject to review by any new entity created to review/certify standards and assessments?

	Total	Yes	No
Schools/Districts	21	1	5
States	14	1	4
Associations	22	0	4
Higher Education	18	1	1
Measurement	12	1	5
Business	6	0	1
Individuals	9	1	1
Other Government	2	0	0
Technical Assist	3	0	1
	107	5	22

5. Achievement Levels for the National Assessment

Most of those who responded indicated their belief that reporting NAEP results using achievement levels can improve public understanding and should be continued. The position of EIAC is that achievement levels be continued because of their potential to increase public understanding of NAEP results. The discussion of the Forum participants was generally consistent with this view, but included the suggestion that work continue in improving the process by which achievement levels are set. In citing the usefulness of achievement levels, one respondent described how its state charted performance of students against achievement levels and, as a result, is reviewing the mathematics programs to which these students are exposed.

Some respondents noted that technical problems accompanied the initial effort to set achievement levels and declined to respond directly until "utility, reliability and validity are confirmed." A few respondents suggested that the particular approach taken (i.e. modified Angoff) is inappropriate. One respondent suggested that "Achievement levels should be driven by curricular and national social considerations that are informed by, but independent of specific assessment techniques or test questions." Another respondent voiced the concern that "NAEP might come to pre-empt state and local efforts to set standards..."

Table 5.

Do achievement levels improve (or have the potential of improving) public understanding of National Assessment results?

	Total	Yes	No
Schools/Districts	21	5	2
States	14	6	2
Associations	22	2	2
Higher Education	18	5	2
Measurement	12	5	1
Business	6	3	0
Individuals	9	1	0
Other Government	2	1	0
Technical Assist	3	1	0
	107	29	9

6. An International Component to NAEP

Most who responded thought that it would be useful to link international results with NAEP. EIAC stated that (1) international comparisons can be informative and should be included in NAEP; and (2) expectations and performance of students from other countries should be taken into account in establishing NAEP frameworks and achievement levels. The discussion of the Forum was mixed--some participants suggested that linking NAEP with international assessments could be useful if the tests measure the same things, are equally reliable, and given to comparable populations; other participants felt that administering NAEP items to students in other countries would be a waste of their time.

Respondents in favor of an international component in NAEP expressed the view that global competition demands that the U.S. give some, but not exclusive, consideration to the curricula and student performance of other nations in deciding what is important for U.S. students to learn. As one respondent put it: "One of the pressing issues of the 1990's is the ability of the U.S. to compete successfully in a global economy...An international comparison would prove a check and balance and help us find the best of all curricular approaches." Another said that "Foreign experience is essential to the identification and setting of reasonable standards." The U.S. Board on International Comparative Studies in Education agreed that "international benchmarks on NAEP scales could...enrich the interpretation of NAEP scales and help guide the definition of meaningful achievement levels..." but also warned of the technical and procedural pitfalls to avoid.

Other respondents were concerned that international comparisons are technically suspect and difficult to achieve. "Comparison of student performance among different nations is problematic, since the educational systems differ dramatically in terms of goals, resources, etc." "Schooling in other nations is so different that it may be inappropriate to offer comparisons or reference points as a way of understanding U.S. student performance." "Such an effort may well distract from [NAEP's mission]."

Table 6.

Would it be useful to see how students in other nations perform on the National Assessment or on certain NAEP items as a reference point for understanding U.S. student performance?

	Total	Yes	No
Schools/Districts	21	5	2
States	14	4	1
Associations	22	3	2
Higher Education	18	5	2
Measurement	12	6	3
Business	6	3	1
Individuals	9	1	1
Other Government	2	2	0
Technical Assist	3	1	0
	107	30	12

7. NAEP as an Anchor for Linking State and Local Assessment Systems with National and International Results

Most who responded favored allowing the use of NAEP as an anchor as long as such use was voluntary and included appropriate controls to assure the credibility and integrity of NAEP. A few favored such use at the state level only. One respondent stated that it would be an attractive option for states or districts, but cautioned that "...it would be premature without further R&D." Another pointed out that "Technically, one cannot assume that...separately developed local tests that have each been equated with NAEP are then likewise equated with each other."

EIAC's position is that states and districts be allowed to use NAEP as an anchor, that the federal government provide resources for research and development for such uses, and that the federal government provide oversight of any eventual such uses. Most of the respondents who favored NAEP as an anchor also felt that research and development should be supported by the federal government and that the federal government had a right and an obligation to provide oversight. The Forum participants expressed doubts that there is a need to anchor state or local data, questioned the validity of using NAEP for such purposes, and felt that pursuing this course would undermine NAEP as a monitor.

Others expressed the view that the "technical difficulties are numerous and not easily overcome" and a concern that such uses of NAEP would "deflect the contractor staff from the continued improvement of the assessment itself and the interpretation of results..."

Table 7.

Should states and local districts be permitted to use the National Assessment as an anchor test for comparability purposes?

	Total	Yes	No
Schools/Districts	21	8	3
States	14	7	0
Associations	22	5	1
Higher Education	18	1	2
Measurement	12	5	4
Business	6	3	0
Individuals	9	1	0
Other Government	2	1	0
Technical Assist	3	0	1
	107	31	9

8. Removing the Prohibition against Using National Assessment Results at the District or School Level

Most who responded answered affirmatively to the question in Table 8. below. They cited NAEP's quality, credibility and extensive background information and their belief that "...the use of assessment data from NAEP instruments [would be] extremely valuable to local schools and districts..." and would provide valuable "...information...regarding progress and problems of their students." A theme from state-based respondents was that data are more instructionally useful when disaggregated below the state level, that sample size in some schools and districts is large enough under current procedures to support reasonable estimates of performance, that receiving such reports would make the assessment more worthwhile for the participants and that the withholding of such information by the federal government seems inappropriate.

EIAC adopted the position that the prohibition should be continued with respect to state-level assessments, but that consideration should be given to permitting below-state reporting at local cost and option for other, national-only assessments. Most of the Forum participants expressed the view that the prohibition should not be lifted, primarily because of concerns that it inappropriately would transform NAEP into a high-stakes test with attendant negative consequences (e.g. teaching to the test and distortions of test results). However, a few Forum participants felt that the prohibition should be lifted, because of NAEP's quality and because it is a "public" resource. One pointed out that the student enrollment in New York city exceeds that of 31 states, and that at least a pilot project for large districts should be considered.

Some respondents questioned whether NAEP is an appropriate instrument for below-state reporting, especially for diagnosing in-classroom instructional performance. Others worried that "local districts [would] be forced into [participating] by eager state bureaucrats..." or that "...high-stakes applications may be tied to NAEP by [state] legislature[s]."

Table 8.

Should states and districts be permitted to use NAEP-related test items, instruments and analytic mechanisms, at their discretion and cost, and according to NAEP requirements, to report local (but not individual student) results?

	Total	Yes	No
Schools/Districts	21	8	2
States	14	7	2
Associations	22	7	2
Higher Education	18	3	3
Measurement	12	4	6
Business	6	3	0
Individuals	9	1	1
Other Government	2	1	0
Technical Assist	3	0	1
	107	34	17

9. Annual Assessments and Reporting

Slightly more than half (about 55%) of those who responded are in favor of permitting annual assessments. EIAC favors annual, voluntary state-level assessments, limited to two subjects in three grades per year, with participation in selected subjects and grades at the discretion of the state, and with subjects tested no more frequently than once in four years. Most of the Forum participants agreed with the idea of annual assessments; however, a few felt strongly that it would result in an excessive burden on schools and could act as a powerful disincentive to participation.

Those in favor felt that an annual assessment schedule would keep the number of assessments per year manageable, allow for planning and continuity within states, and permit periodic assessment of additional subjects. Those who favored the biennial schedule felt that "[t]here is already too much...testing burden on American kids," "[t]he need for an annual assessment is not clear," and it would be too costly.

Table 9.

Should the National Assessment legislation be amended to permit annual assessments?

	Total	Yes	No
Schools/Districts	21	2	4
States	14	5	2
Associations	22	4	1
Higher Education	18	2	2
Measurement	12	3	4
Business	6	2	0
Individuals	9	0	1
Other Government	2	0	0
Technical Assist	3	1	0
	107	19	14

10. Additional Issues

Several respondents identified issues that are not addressed in the discussion paper on the Future of NAEP.

a. Testing Linguistic Minorities

Several respondents and a Forum participant raised the concern that the significant growth in the numbers of limited English proficient students nationally, and the large proportion of such students in certain states, poses technical and policy matters that need to be addressed by the National Assessment. For example, one respondent estimated that "...by the year 2000, the majority of students in California may well be nonnative English speakers." The question is--how representative can state/national NAEP results be if such students are not included in the assessment; and how can such students be fairly assessed?

b. Participation of Disabled Students

Two respondents raised the concern that exclusion rates from NAEP of disabled students are very high and too variable among the states. The point made by the respondents is that high and variable, unexplained exclusions of disabled students impact on the comparability of state assessments and the comprehensiveness of national assessments.

c. Linking Curricular Practices and Policies to Outcomes

Two respondents addressed the potential in NAEP, through the background questionnaires, to do a better job of collecting and analyzing information on the conditions of instruction in the nation's classrooms and linking this with information on student outcomes. One of these respondents suggested using NAEP as a mechanism to define and collect information on the educational delivery standards discussed in the report of the National Council on Education Standards and Tests.

Prepared: November, 1992; Ad Hoc Committee on the Future of NAEP
Members: Jerry Hume (Chair), Honorable Evan Bayh, Mary Blanton, Michael Glode, Christine Johnson, Mark Musick, Thomas Topuzes.
Staff: Ray Fields

THE FUTURE OF THE NATIONAL ASSESSMENT OF EDUCATIONAL PROGRESS (NAEP)

A Resolution by the Education Information Advisory Committee (EIAC) Assessment Subcommittee -- September 17, 1992 -- Alexandria, Virginia

Following review and detailed discussion of "The Future of the National Assessment of Educational Progress (NAEP) Discussion Paper" prepared by the National Assessment Governing Board (NAGB), it is recommended that:

(1) the contents of NAEP assessments be based on both national standards and current as well as evolving instructional practices;

(2) the NAEP assessment frameworks be developed to incrementally lead instruction;

(3) the NAEP legislation be amended to allow for annual assessments and to continue to conduct state-level assessments, on a voluntary basis;

(4) the state assessment component of NAEP be limited to two (2) subjects areas at three (3) grades per year to limit respondent burden;

(5) the volunteer state participants be allowed to pick and choose their specific level of participation from the subject areas and grades to be assessed during a given year;

(6) the NAEP state assessment component subject area assessments be placed on four-year assessment cycles;

(7) the NAEP state assessment component include, but not be limited to:

Reading,
Writing,
Mathematics,
Science,
Social Studies (including U.S. History and Geography),
Foreign Languages,
The Arts,
Literature, and
SCANS competencies;

(8) NAEP continue to measure both long- and short-term student achievement trends, but that measurement of long-term student achievement trends be discontinued when the resulting data provides misleading information regarding student achievement resulting from out-dated items in relation to assessment frameworks and national standards;

(9) state and local districts be allowed to use NAEP as an anchor test for comparability purposes, with the federal government providing resources to research and develop such uses of NAEP, and the federal government serving an oversight role in such uses of NAEP;

(10) the prohibition on NAEP preventing district- and school-level results remain in effect for any subject areas involved with the state assessment component, though proposals for district- and school-level results should be considered in other assessment areas;

(11) achievement levels continue to be used for reporting because of their potential to increase public understanding of the results, but with work to improve the achievement levels setting process being continued;

(12) international comparisons have informative value and should be included in NAEP, with student expectations and performance from other countries being taken into account in establishing NAEP frameworks and achievement levels;

(13) until the proposed National Education Standards and Assessment Council is statutorily defined and operationalized, the relationship of NAEP to the proposed Council cannot be commented upon; and

(14) the costs of the NAEP state assessment component be shared between the federal government and the volunteer states.

TAB 3

THE NATION'S REPORT CARD

National Assessment Governing Board

National Assessment of Educational Progress

Alternative Future Scenarios

Shortly after the Education Summit in Charlottesville in 1989, Deputy Secretary Ted Sanders asked the Governing Board to prepare a paper that would discuss the implications of the six national education goals for the National Assessment. The paper was prepared and in December 1989 the Governing Board unanimously adopted a document entitled "Positions on the Future of NAEP."

The current examination of NAEP's role was prompted by several successive, related events:

- decision by the National Education Goals Panel that NAEP should be the primary source of data on national and state progress toward goal 3;
- recommendation of the National Council on Education Standards and Tests (the National Council) for the development of national content and student performance standards, and a national system of assessments;
- grant awards made by the Department of Education that will result in the development of voluntary national content and performance standards in English, science, history, geography, the arts and civics.

STANDARDS AND ASSESSMENTS

Work has begun on national content and performance standards. They are expected to be ready for dissemination during the 1993-1994 school year. Standards for mathematics were released by the National Council of Teachers of Mathematics in 1989. There appears to be a general consensus, among educators and in Congress, that developing such standards is a worthwhile endeavor.

There is considerably less consensus about whether and how to proceed toward a national system of assessments. The National Council recommended a two-tier approach:

- individual student assessments, developed by states, districts, commercial publishers, and others to measure student performance against the national content and performance standards;
- assessments of samples of students by NAEP, appropriately aligned with national standards as they are developed, to monitor the nation's and the states' progress toward goals 3 and 4 of the National Education Goals.

800 North Capitol Street, N.W.
Suite 825
Mailstop 7583
Washington, D.C. 20002-4233
(202) 357-6938

COORDINATING STRUCTURE

The Council also recommended the creation of a coordinating structure that would

- establish guidelines for standards-setting and assessment development;
- establish general criteria to determine appropriateness of standards and assessments;
- certify, jointly with the Goals Panel, content and student performance standards and criteria for assessments as world class.

This coordinating structure would be the National Education Standards and Assessments Council (NESAC).

The 102nd Congress considered legislation related to the Goals Panel and the Council recommendations. Although legislation was reported out of conference, none was enacted. The results of the conference indicate that, while there is palpable support for the development of content standards, policy on how to proceed in developing a national system of assessments is less clear.

With respect to standards, the conference report gave NESAC somewhat broad authority:

- develop criteria for reviewing standards;
- make recommendations to the Goals Panel on the selection of awardees for grants to develop content and delivery standards;
- review, and make recommendations to the Goals Panel regarding, content standards.

It is important to note that, with respect to standards, incentives are provided to prompt their development. That is, the conference report provided resources for the development of standards and a system for their review.

With respect to assessments, the conference report gave NESAC much more limited authority:

- develop criteria for the development of different types of model math and science assessments, to be developed with funds under the Eisenhower Act and to be tied to national math and science standards;

- o make recommendations to the Goals Panel on the selection of awardees for grants under the Eisenhower Act to develop model science and mathematics assessments.

The conference report does not give NESAC the general authority for developing criteria for and reviewing assessments that was envisioned by the Council. Thus, many questions remain about what a national system of assessments would look like, how separately developed assessments would fit together, whether mechanisms would be used to permit comparability across students, schools, districts and states, and what incentives would be employed to prompt any change from the status quo.

HOW WILL THE EDUCATION SYSTEM BEHAVE?

Similarly, it is unknown what the behavior of states and schools would be in response to the Council recommendations or to the policies indicated in the conference report. The specific dimensions of a national system of assessments remain undefined; thus, in envisioning NAEP's future role, one can only speculate based on a set of plausible scenarios. These scenarios are described below and take into account the following set of assumptions:

1. National content and performance standards for English, science, history, geography, the arts, and civics will not be disseminated before school year 1993-94 (mathematics standards were released by NCTM in 1989).

2. Some organizations developing content standards may not develop student performance standards.

3. Being voluntary, national standards will not be adopted universally by states, districts and schools, public and private, upon release.

4. Even if universally adopted, the rate and quality of implementation will vary among schools within districts, among districts within states, and among states.

5. Something approaching "full" implementation of national content standards will not occur sooner than 5-10 years after their release, during which time they may be revised or updated.

6. National content and performance standards will have an effect on the products of textbook and commercial test publishers.

7. The locus of control over education policy will continue to reside within states, local school boards and private schools.

8. NAEP has a continuing role to play in monitoring education progress.

SCENARIO ONE: FULL ADOPTION OF THE COUNCIL PROPOSAL

The Council proposal provides for the continuation of the Goals Panel and the establishment of NESAC, appointed by the Goals Panel. The Goals Panel and NESAC jointly would be responsible for certifying content and student performance standards and criteria for assessments.

In this scenario, NESAC is established during 1993. The appointment, funding and staffing process is completed within four months. Certification guidelines, criteria and procedures are completed in another twelve months, i.e. during the 1994-95 school year. The NCTM mathematics standards are used as a test case during this developmental period, since they already exist. By this time, NESAC has before it the proposals for standards in English, science, history, geography, the arts, and civics, all recently completed, and perhaps one or two additional proposals as well.

The review process is rigorous, and NESAC takes not less than 3-4 months to develop recommendations to the Goals Panel. (It is possible that the sudden workload would require more time.) The Goals Panel review of the NESAC recommendations takes no less than another 1-2 months. It is likely that content standards would not be certified in sufficient time prior to the 1995-96 school year to permit state and local adoption, dissemination and training for implementation in 1995-96. Instead, 1995-96 is a period when states, districts and schools begin to examine the new standards, make decisions about whether to adopt the standards in part or in whole, and plan for implementation.

At the same time, textbook and test publishers are developing new products that are ready for distribution in about 12-18 months after standards are certified. Assuming that there are no major delays, systems for instructional support and assessment would begin to come on line during the 1996-97 school year. Earliest, and probably somewhat limited, state and local assessment results would be available toward the end of 1997. These initial results are considered "field trials" to be repeated again in 1998.

State policy on adoption of the standards and on assessment is not monolithic. Some states adopt each of the certified standards as guidelines for curriculum. A subset of these states are required by their legislatures to develop state assessment programs to measure the performance of individual students against the standards to be used in making promotion and graduation decisions and decide to employ a census model in selected grades. Generally, testing occurs prior to transition grades so that students have multiple opportunities to receive instruction that will enable them to demonstrate adequate competency. Other states pick and choose among the standards they embrace. These states decide that the purpose of assessment is to assess school and district

implementation of the standards and, therefore, adopt a matrix sample approach to testing. Because of varying state requirements, school grade structures, and precedent, student testing across states does not occur in identical grades or at identical times. Some states decide to pool efforts and develop common tests and reporting procedures, other states go it alone.

IMPLICATIONS

Under the best of circumstances, the chance is low, perhaps 20-30 percent, that, before 1998, a full-blown "national system of assessments" tied to national standards and providing comparable results will be in place. It remains particularly unclear how what is envisioned by NESAC would necessarily lead to something that meets the definition of the term "system" (i.e. an assemblage of parts forming a unitary whole). This is particularly true because there is no clear mechanism yet defined that would link the various assessments that are anticipated; the standards are a necessary but not sufficient condition for such a mechanism, and it seems unlikely that they will be adopted by all parties simultaneously nor implemented with similar care and intensity. Likewise, state and local policy on when and what to assess is highly likely to remain variable through the end of the decade.

In this scenario, NAEP would not be impeded from fulfilling the role envisioned for it by the Council. NAEP could continue to collect and report national data on student performance. Unless the Goals Panel and the states decide that state-comparable data are not important for tracking progress toward national education goals and useful for informing policy, NAEP could continue to collect state-level data as well. NAEP would continue the practice begun with the NCTM standards--using the national consensus process employed in developing NAEP frameworks to advise on the inclusion of content standards as they are being developed and, through successive administrations of a subject area assessment, appropriately adjust the framework and/or test specifications. At present, support for this approach is overwhelming. (See analysis of responses to discussion paper on the future of NAEP). In addition, through the end of the decade, NAEP could promote and participate in research and development on mechanisms for comparing separate assessments.

SCENARIO 2: CONTENT STANDARDS DEVELOPED IN THE ABSENCE OF A CERTIFICATION PROCESS

In this scenario, the standards projects now underway are completed according to the schedule in scenario 1. Development of standards in other subjects may be completed by interested organizations. But no single, formal process is established for certifying the standards or for reviewing assessments. Instead, the "marketplace of ideas" serves as an informal certification mechanism. Just as the NCTM standards are being reviewed and adopted independently by educators throughout the nation on the basis of their perceived merits, so would the standards developed in other disciplines.

IMPLICATIONS

Without a "seal of approval," the attention and weight given to the standards by the field may be lessened somewhat. Adoption and implementation would not be faster, and could be slower, although at least six or seven months would not be expended for what would have been review and certification by NESAC and the Goals Panel. However, to the extent that the standards do influence policy and instruction, the organizations that develop the standards may have an interest in assuring that unwarranted claims are not made about the incorporation of the standards in textbooks and tests. These organizations might seek to establish their own "certification" procedures to validate claims made by publishers about their products and by education officials about the performance of their students and schools.

The major consideration for NAEP in this scenario is the possible "inappropriate influence [of] special interests" (GEPA Section 406(i)(5)(C)(i)). That is, the content of each NAEP assessment should be informed by the standards-setting initiatives, but should be determined through the national consensus process required by law.

SCENARIO 3: ADOPTION OF SCHOOL DELIVERY STANDARDS

A point of great contention with respect to the Council proposal is the issue of school delivery standards. School delivery standards are not defined in the proposal, but are described as "developed by the states collectively from which each state could select the criteria that it finds useful for the purpose of assessing a school's capacity and performance." The conference report defines school delivery standards as "the criteria for, and means of assessing, the resources, practices and conditions necessary at each level of the education system (schools, school districts and states) to provide all students with an opportunity to learn the material in...voluntary national content standards..."

The contention arises from concerns on the part of some that the existence of national school delivery standards would lead to prescriptive national requirements for allocating resources to schools; however, the desire to have school delivery standards arises from the concern of others that adopting high standards without the means of assuring fair opportunity for all students to achieve the standards would be unconscionable.

IMPLICATIONS

Some respondents have suggested that NAEP's background questionnaires be revised to collect some of the information that might comprise school delivery standards (see analysis of responses). The Governing Board has previously taken up the question of how NAEP might better report on opportunity to learn. Recent frameworks developed through the national consensus process increasingly have included suggestions for background questions that address the degree to which the assessment objectives are reflected in instruction provided to the students who are assessed. By the same token, it is beyond NAEP's powers to correlate resource allocation inputs and student performance. However, it is clear that as content standards are developed and implemented, NAEP could provide useful policy information at the state and national levels that provide indications of the degree of implementation national standards and the relationship to student achievement.

SCENARIO 4: RAPID RESPONSE BY STATES

In this scenario, states organize themselves during 1993 to put mechanisms in place for a national system of assessments and to prepare for the release of national standards in 1994. Clusters of states form regional curriculum and test development collaboratives; others proceed independently.

The consensus process used by the organizations developing standards incorporates state review of each draft, thus states are well-informed of the shape the standards are taking and do not wait until their official release to begin substantive work on revising curricula and tests.

Upon release of the standards in 1994, nearly or virtually every state legislature provides funds for inservice training of all affected teachers in the state to implement the standards, new curricula and testing. The inservice training begins during the 1993-94 school year and is continued intensively through pre-service training during the summer of 1994. Instruction according to the new standards begins in school year 1994-95. Field trials of individual student tests occur in the months of April and May of 1995. Data is reported in October 1995. States find that their test results are not comparable measures of performance.

IMPLICATIONS

Achievement of this scenario requires a level of resources, singlemindedness of purpose and exercise of political will and administrative acumen never before seen in America's decentralized education system. There is only a small chance that it would be carried out as smoothly as described.

A conservative estimate of the cost of teacher training/planning would be $500 per teacher. To reach the 2.4 million public school teachers, the total cost would be about $1.2 billion. Anecdotal evidence indicates that this would represent perhaps a 20-fold increase in state and local resources for substantive inservice training. In addition, the desire to minimize student test burden and the high profile given to state test activity both would act as disincentives for states to participate in state-level versions of the National Assessment. However, for those that wanted to report state and national results in comparable terms, some individual states and regional clusters might approach the National Assessment with a proposal to conduct equating studies using the National Assessment as a "standard" against which their results would be reported.

Prepared: November, 1992; Ad Hoc Committee on the Future of NAEP
Members: Jerry Hume (Chair), Honorable Evan Bayh, Mary Blanton, Michael Glode, Christine Johnson, Mark Musick, Thomas Topuzes
Staff: Ray Fields

THE NATION'S REPORT CARD

National Assessment Governing Board

National Assessment of Educational Progress

Adopted Unanimously: November 21, 1992

RESOLUTION

Whereas the National Assessment Governing Board is responsible for selecting subject areas to be assessed;

Whereas the Governing Board, after two years of deliberations and with extensive external advice, has identified a set of non-mandated subjects to be assessed;

Whereas the Governing Board believes that the National Assessment should be conducted annually;

Whereas the Governing Board understands that the Third International Mathematics and Science Study will be conducted in 1995 and 1999 and believes that conducting NAEP science and mathematics assessments in those years would facilitate the linking of state and national NAEP results with international results;

Whereas the Governing Board recognizes that other subject areas are important and intends not to foreclose consideration of such subject areas for assessment in the future in the form of full assessments and/or more limited probes; and

Whereas the Governing Board recognizes that legislative authority is needed in order to conduct assessments on a regular annual schedule;

Therefore, the Executive Committee recommends that the National Assessment Governing Board endorse the following schedule of assessments through the year 2000:

1994 -	Reading
	Geography
	U.S. History
1995 -	Mathematics
	Science
	Writing
1996 -	Reading
	Arts
	Civics
1997 -	Mathematics
	Economics
	Foreign Language
1998 -	Reading
	Geography
	U.S. History/World History
1999 -	Mathematics
	Science
	Writing
2000 -	Reading
	Civics
	(Optional Subject)

800 North Capitol Street, N.W.
Suite 825
Mailstop 7583
Washington, D.C. 20002-4233
(202) 357-6938